ON THE SECOND ORDER DERIVATIVES OF CONVEX FUNCTIONS ON THE HEISENBERG GROUP

CRISTIAN E. GUTIÉRREZ AND ANNAMARIA MONTANARI

1. I

A classical result of Aleksandrov asserts that convex functions in $\mathbb{R}^n$ are twice differentiable a.e., and a first step to prove it is to show that these functions have second order distributional derivatives which are measures, see [3, pp. 239-245]. On the Heisenberg group, and more generally in Carnot groups, several notions of convexity have been introduced and compared in [2] and [5], and Ambrosio and Magnani [1, p. 3] ask the natural question if a similar result holds in this setting. Recently, these authors proved in [1, Theorem 3.9] that $BV^2_{\mathbb{H}}$ functions on Carnot groups, that is, functions whose second order horizontal distributional derivatives are measures of H-bounded variation, have second order horizontal derivatives a.e., see Subsection 2.1 below for precise statements and definitions. On the other hand and also recently, Lu, Manfredi and Stroffolini proved that if u is an $\mathcal{H}$–convex function in an open set of the Heisenberg group $\mathbb{H}^1$ in the sense of the Definition 2.4 below, then the second order symmetric derivatives $(X_iX_ju + X_jX_iu)/2$, $i, j = 1, 2$, are Radon measures [5, Theorem 4.2], where X_j are the Heisenberg vector fields defined by (2.1). Their proof is an adaptation of the Euclidean one, it is based on the Riesz representation theorem, and it can be carried out in the same way for $\mathbb{H}^n$. However, to prove that $\mathcal{H}$–convex functions u are $BV^2_{\mathbb{H}}$, one should show that the non symmetric derivatives X_iX_ju are Radon measures. Since the symmetry of the horizontal derivatives is essential in the proof of [5, Theorem 4.2], this prevents these authors to answer the question of whether or not the class of $\mathcal{H}$–convex functions is contained in $BV^2_{\mathbb{H}}$.

The purpose in this paper is to establish the existence a.e. of second order horizontal derivatives for the class of $\mathcal{H}$-convex functions in the sense of Definition 2.4. We will actually prove the stronger result that every $\mathcal{H}$-convex function belongs to the class $BV^2_{\mathbb{H}}$ answering the question posed by Ambrosio and Magnani in the setting of the Heisenberg group. In order to do this we use the technique from our work [4] which we shall briefly explain. Indeed, following an approach recently used by Trudinger and Wang to study Hessian equations [7], we proved in [4] integral estimates in $\mathbb{H}^1$ in terms of the following Monge–Ampère type operator: $\det \mathcal{H}(u)+12\,(u_t)^2$, see Definition 2.4. We first established, by means of integration by parts, a comparison principle for smooth functions, and then

February 1, 2008

The first author was partially supported by NSF grant DMS–0300004, and thanks the University of Bologna and the INDAM for the kind hospitality and their support. The second author was partially supported by the University of Bologna, funds for selected research topics.

extended this principle to "cones". Together with the geometry in $\mathbb{H}^1$, this leads to an Aleksandrov type maximum principle [4, Theorem 5.5]. Moreover, in [4, Proposition 6.2] we proved an the estimate of the oscillation of $\mathcal{H}$–convex functions. This estimate furnishes L^2 estimates of the Lie bracket $[X_1, X_2]u = -4\partial_t u$ of $\mathcal{H}$–convex functions on $\mathbb{H}^1$ and permits to fill the gap between the results in [5, Theorem 4.2] and [1, Theorem 3.9], and to prove that

$$X_i X_j u = \frac{[X_i, X_j]u}{2} + \frac{(X_i X_j + X_j X_i)u}{2}, \quad i, j = 1, 2,$$

are Radon measures.

Following the route just described in $\mathbb{H}^1$, in this paper we introduce in $\mathbb{H}^n$ the operator $\sigma_2(\mathcal{H}(u)) + 12nu_t^2$, where σ_2 is the second elementary symmetric function of the eigenvalues of the matrix $\mathcal{H}(u)$, we define the notion of $\sigma_2(\mathcal{H})$–convex function related to this operator, and as a main tool we establish a comparison principle for $\sigma_2(\mathcal{H})$–convex functions, see Definition 2.5 and Theorem 3.1. In this frame, we next establish an oscillation estimate, Proposition 4.2, which yields as a byproduct L^2 estimates of $\partial_t u$ in $\mathbb{H}^n$ for a class of functions bigger than the class of $\mathcal{H}$–convex functions. We apply these estimates to obtain that the class of $\mathcal{H}$-convex functions is contained in $BV^2_{\mathbb{H}}$, and as a corollary of [1, Theorem 3.9] it follows that $\mathcal{H}$-convex functions have horizontal second derivatives a.e.

The paper is organized as follows. Section 2 contains preliminaries about $\mathbb{H}^n$, $BV_{\mathbb{H}}$ functions, and the definitions of $\mathcal{H}$–convexity and $\sigma_2(\mathcal{H})$–convexity. In Section 3 we prove a comparison principle for C^2 functions. Section 4 contains the oscillation estimate and the construction of the analogue Monge–Ampère measures for $\sigma_2(\mathcal{H})$–convex functions. Finally, in Section 5 we prove Aleksandrov's type differentiability theorem for $\mathcal{H}$–convex functions in $\mathbb{H}^n$.

Acknowledgements. We are greatly indebted to Bianca Stroffolini for some useful discussions.

At the Workshop on Second Order Subelliptic Equations and Applications, Cortona June 2003, we learnt from Nicola Garofalo that in a joint paper with Federico Tournier, they extended to higher dimensions the Aleksandrov type maximum principle proved by us in [4, Theorem 5.5] for $\mathcal{H}$–convex functions in $\mathbb{H}^1$.

2. P , $\mathcal{H}$– $\sigma_2(\mathcal{H})$–

Let $\xi = (x, y, t), \xi_0 = (x_0, y_0, t_0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$, and if $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, then $x \cdot y = \sum_{j=1}^n x_j y_j$. The Lie algebra of $\mathbb{H}^n$ is spanned by the left-invariant vector fields

$$X_j = \partial_{x_j} + 2y_j\,\partial_t, \quad X_{n+j} = \partial_{y_j} - 2x_j\,\partial_t \quad \text{for } j = 1, \dots n. \tag{2.1}$$

We have $[X_j, X_{n+j}] = X_j X_{n+j} - X_{n+j} X_j = -4\partial_t$ for every $j = 1, \dots n$, and $[X_j, X_i] = X_j X_i - X_i X_j = 0$ for every $i \neq n + j$. If $\xi_0 = (x_0, y_0, t_0)$, then the non–commutative multiplication law in $\mathbb{H}^n$ is given by

$$\xi_0 \circ \xi = (x_0 + x, y_0 + y, t_0 + t + 2(x \cdot y_0 - y \cdot x_0)),$$

and we have $\xi^{-1} = -\xi$, $(\xi_0 \circ \xi)^{-1} = \xi^{-1} \circ \xi_0^{-1}$. In $\mathbb{H}^n$ we define the gauge function

$$\rho(\xi) = \left((|x|^2 + |y|^2)^2 + t^2\right)^{1/4},$$

and the distance

$$d(\xi, \xi_0) = \rho(\xi_0^{-1} \circ \xi). \tag{2.2}$$

The group $\mathbb{H}^n$ has a family of dilations that are the group homomorphisms, given by

$$\delta_\lambda(\xi) = (\lambda x, \lambda y, \lambda^2 t)$$

for $\lambda > 0$. Then

$$d(\delta_\lambda \xi, \delta_\lambda \xi_0) = \lambda\, d(\xi, \xi_0).$$

For more details about $\mathbb{H}^n$ see [6, Chapters XII and XIII].

2.1. *$BV_{\mathbb{H}}$* **functions.** For convenience of the reader, we collect here some definitions and a result from Ambrosio and Magnani [1] particularized to the Heisenberg group that will be used in the proof of Theorem 5.1.

We identify the vector field X_j with the vector $(e_j, \overrightarrow{0}, 2y_j)$ in $\mathbb{R}^{2n+1}$ for $j = 1, \cdots, n$, and with the vector $(\overrightarrow{0}, e_j, -2x_j)$ for $j = n+1, \cdots, 2n$. Here e_j is the jth-coordinate vector in $\mathbb{R}^n$ and $\overrightarrow{0}$ is the zero vector in $\mathbb{R}^n$. Given $\xi = (x, y, t) \in \mathbb{R}^{2n+1}$, with this identification we let $\{X_j(\xi)\}_{j=1}^{2n}$ be the vectors with origin at ξ and set $H_\xi = \operatorname{span}\{X_j(\xi)\}$. The set H_ξ is a hyperplane in $\mathbb{R}^{2n+1}$. Given $\Omega \subset \mathbb{R}^{2n+1}$ we set $H\Omega = \cup_{\xi \in \Omega} H_\xi$. Consider $\mathcal{T}_{c,1}(H\Omega)$ the class functions $\phi : \Omega \to \mathbb{R}^{2n+1}$, $\phi = \sum_{j=1}^{2n} \phi_j X_j$ that are smooth and with compact support contained in Ω and denote by $\|\phi\| = \sup_{\xi \in \Omega} \sum_{j=1}^{2n} |\phi_j(\xi)|$.

Definition 2.1. *We say that the function $u \in L^1(\Omega)$ is of H-bounded variation if*

$$\sup\left\{\int_\Omega u \operatorname{div}_X \phi \, dx : \phi \in \mathcal{T}_{c,1}(H\Omega),\ \|\phi\| \le 1\right\} < \infty,$$

where $\operatorname{div}_X \phi = \sum_{i=1}^{2n} X_i \phi_i$. The class of these functions is denoted by $BV_{\mathbb{H}}(\Omega)$.

Definition 2.2. *Let $k \ge 2$. The function $u : \Omega \to \mathbb{R}$ has H-bounded k variation if the distributional derivatives $X_j u$, $j = 1, \cdots, 2n$ are representable by functions of H-bounded $k-1$ variation. If $k = 1$, then u has H-bounded 1 variation if u is of H-bounded variation. The class of functions with H-bounded k variation is denoted by $BV_{\mathbb{H}}^k(\Omega)$.*

Theorem 2.3 (Ambrosio and Magnani [1], Theorem 3.9)**.** *If $u \in BV_{\mathbb{H}}^2(\Omega)$, then for a.e. ξ_0 in Ω there exists a polynomial $P_{[\xi_0]}(\xi)$ with homogeneous degree ≤ 2 such that*

$$\lim_{r \to 0^+} \frac{1}{r^2} \fint_{U_{\xi_0,r}} |u(\xi) - P_{[\xi_0]}(\xi)|\, d\xi = 0,$$

where $U_{\xi_0,r}$ is the ball centered at ξ_0 with radius r in the metric generated by the vector fields X_j, and

$$P_{[\xi_0]}(\xi) = P_{[\xi_0]}\left(\exp\left(\sum_{j=1}^{2n} \eta_j X_j + \eta_{2n+1}[X_1, X_2]\right)(\xi_0)\right) = \sum_{|\alpha| \le 2} c_\alpha \eta^\alpha,$$

with $\alpha = (\alpha_1, \dots, \alpha_{2n+1})$, $c_\alpha \in \mathbb{R}$, $\eta^\alpha = \eta_1^{\alpha_1} \cdots \eta_{2n+1}^{\alpha_{2n+1}}$ *and* $|\alpha| = \sum_{j=1}^{2n} \alpha_j + 2\alpha_{2n+1}$. *

2.2. **$\mathcal{H}$–convexity and $\sigma_2(\mathcal{H})$–convexity.** For a C^2 function u, let $X^2 u$ denote the non symmetric matrix $[X_i X_j u]$. Given $c \in \mathbb{C}$ and $u \in C^2(\Omega)$, let

$$\mathcal{H}_c(u) = X^2 u + c u_t \begin{bmatrix} 0_n & I_n \\ -I_n & O_n \end{bmatrix}.$$

Definition 2.4. *The function $u \in C^2(\Omega)$ is $\mathcal{H}$–convex in Ω if the $2n \times 2n$ symmetric matrix*

$$\mathcal{H}(u) = \mathcal{H}_2(u) = \left[\frac{X_i X_j u + X_j X_i u}{2} \right]$$

is positive semidefinite in Ω.

Notice that the matrix $\mathcal{H}_c(u)$ is symmetric if and only if $c = 2$. Also, if $\langle \mathcal{H}_c(u)\xi, \xi \rangle \geq 0$ for all $\xi \in \mathbb{R}^{2n}$ and for some c, then this quadratic form is nonnegative for all values of $c \in \mathbb{R}$.

Definition 2.5. *The function $u \in C^2(\Omega)$ is $\sigma_2(\mathcal{H})$–convex in Ω if*

(1) *the trace of the symmetric matrix $\mathcal{H}(u)$ is non negative,*
(2) *the second elementary symmetric function in the eigenvalues of $\mathcal{H}(u)$*

$$\sigma_2(\mathcal{H}(u)) = \sum_{i<j} \left\{ X_i^2 u X_j^2 u - \left(\frac{X_i X_j u + X_j X_i u}{2} \right)^2 \right\}$$

is non negative.

We extend the definition of $\sigma_2(\mathcal{H})$–convexity to continuous functions.

Definition 2.6. *The function $u \in C(\Omega)$ is $\sigma_2(\mathcal{H})$–convex in Ω if there exists a sequence $u_k \in C^2(\Omega)$ of $\sigma_2(\mathcal{H})$–convex functions in Ω such that $u_k \to u$ uniformly on compact subsets of Ω.*

Remark 2.7. If u is $\mathcal{H}$–convex, then it is $\sigma_2(\mathcal{H})$–convex. The two definitions are equivalent in $\mathbb{H}^1$. Moreover, from [2, Theorem 5.11] we have that if u is convex in the standard sense, then u is $\mathcal{H}$–convex. However, the gauge function $\rho(x, y, t) = \left((|x|^2 + |y|^2)^2 + t^2 \right)^{1/4}$ is $\mathcal{H}$–convex but is not convex in the standard sense.

3. C P

A crucial step in the proof of Aleksandrov's type theorem, Theorem 5.1, is the following comparison principle for C^2 and $\sigma_2(\mathcal{H})$–convex functions.

Theorem 3.1. *Let $u, v \in C^2(\overline{\Omega})$ such that $u + v$ is $\sigma_2(\mathcal{H})$–convex in Ω satisfying $v = u$ on $\partial\Omega$ and $v < u$ in Ω. Then*

$$\int_\Omega \left\{ \sigma_2(\mathcal{H}(u)) + 12\, n\, (\partial_t u)^2 \right\} dz \leq \int_\Omega \left\{ \sigma_2(\mathcal{H}(v)) + 12\, n\, (\partial_t v)^2 \right\} dz,$$

*We can explicitly compute $\eta = (x - x_0, y - y_0, (t_0 - t + 2(x \cdot y_0 - y \cdot x_0))/4)$ by solving the ODE $\xi = \exp\left(\sum_{j=1}^{2n} \eta_j X_j + \eta_{2n+1} [X_1, X_2] \right) (\xi_0)$.

and

$$\int_\Omega \operatorname{trace}\mathcal{H}(u)\,dz \le \int_\Omega \operatorname{trace}\mathcal{H}(v)\,dz.$$

Proof. By arguing as in [4], set

$$S(u) = \sigma_2(\mathcal{H}(u)) = \sum_{i<j}\left\{X_i^2uX_j^2u - \left(\frac{X_iX_ju + X_jX_iu}{2}\right)^2\right\}.$$

We have, by putting $r_{ij} = \dfrac{X_iX_ju + X_jX_iu}{2}$,

$$\frac{\partial S(u)}{\partial r_{ii}} = \sum_{j\neq i} X_j^2 u; \qquad \frac{\partial S(u)}{\partial r_{ij}} = -\left(\frac{X_iX_j + X_jX_i}{2}\right)u, \tag{3.3}$$

and it is a standard fact that if u is $\sigma_2(\mathcal{H})$–convex, then the matrix $\dfrac{\partial S(u)}{\partial r_{ij}}$ is non negative definite, see Section 6 for a proof. Let $0 \le s \le 1$ and $\varphi(s) = S(v + sw)$, $w = u - v$. Then

$$\begin{aligned}
&\int_\Omega \{S(u) - S(v)\}\,dz \\
&= \int_0^1 \int_\Omega \varphi'(s)\,dzds \\
&= \int_0^1 \int_\Omega \left\{\sum_{i,j=1}^{2n} \frac{\partial S}{\partial r_{ij}}(v + sw)\,(X_iX_j)w\right\}\,dzds \\
&= \int_0^1 \int_\Omega \left\{\sum_{i,j=1}^{2n} X_i\left(\frac{\partial S}{\partial r_{ij}}(v + sw)\,X_jw\right) - X_i\left(\frac{\partial S}{\partial r_{ij}}(v + sw)\right)X_jw\right\}\,dzds \\
&= A - B.
\end{aligned}$$

Since $w = 0$ on $\partial\Omega$, $w > 0$ in Ω, then the normal to $\partial\Omega$ is $\nu_X = -\dfrac{Xw}{|Dw|}$. Integrating by parts A we have

$$\begin{aligned}
A &= \int_0^1 \int_\Omega \sum_{i,j=1}^{2n} X_i\left(\frac{\partial S}{\partial r_{ij}}(v + sw)\right)X_jw\,dzds \\
&= \int_0^1 \int_{\partial\Omega} \sum_{i,j=1}^{2n}\left(\frac{\partial S}{\partial r_{ij}}(v + sw)\right)X_jw\,\nu_{X_i}d\sigma(z)ds \\
&= -\int_0^1 \int_{\partial\Omega} \sum_{i,j=1}^{2n}\left(\frac{\partial S}{\partial r_{ij}}(v + sw)\,X_jw\right)\frac{X_iw}{|Dw|}d\sigma(z)ds \\
&= -\frac{1}{2}\int_{\partial\Omega} \sum_{i,j=1}^{2n}\left(\frac{\partial S}{\partial r_{ij}}(u + v)\,X_jw\right)\frac{X_iw}{|Dw|}d\sigma(z) \le 0.
\end{aligned}$$

We now calculate B. Let us remark that for any fixed $j = 1, \dots, 2n$ by (3.3) we have

$$
\begin{aligned}
\sum_{i=1}^{2n} X_i\left(\frac{\partial S}{\partial r_{ij}}\omega\right) &= X_j\left(\frac{\partial S}{\partial r_{jj}}\omega\right) + \sum_{i\neq j} X_i\left(\frac{\partial S}{\partial r_{ij}}\omega\right)\\
&= X_j\left(\sum_{k\neq j} X_k^2\omega\right) - \sum_{i\neq j} X_i\left(\frac{X_iX_j\omega + X_jX_i\omega}{2}\right)\\
&= \sum_{i\neq j}\left(X_jX_i^2\omega - X_i\left(\frac{X_iX_j\omega + X_jX_i\omega}{2}\right)\right)\\
&= \sum_{i\neq j}\left(\frac{[X_j, X_i]X_i\omega}{2} + \frac{[X_j, X_i]X_i\omega}{2} + \frac{X_i[X_j, X_i]\omega}{2}\right)\\
&= 3\sum_{i\neq j}\left(\frac{X_i[X_j, X_i]\omega}{2}\right)\\
&= \frac{3}{2}\begin{cases} X_{j+n}[X_j, X_{j+n}]\omega, & \text{if } j \leq n\\ X_{j-n}[X_j, X_{j-n}]\omega, & \text{if } j > n,\end{cases}
\end{aligned}
$$

where, in the last two equalities, we have used the remarkable fact that $[X_i, [X_j, X_k]] = 0$ for every $i, j, k = 1, \dots, 2n$, and $[X_j, X_i] \neq 0$ iff $i = j \pm n$. Hence,

$$
\begin{aligned}
B &= \int_0^1 \int_\Omega \sum_{i,j=1}^{2n} X_i \left(\frac{\partial S}{\partial r_{ij}}(v+sw) \right) X_j w \, dz ds \\
&= \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} X_{j+n} [X_j, X_{j+n}](v+sw) X_j w \, dz ds \\
&\qquad + \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=n+1}^{2n} X_{j-n} [X_j, X_{j-n}](v+sw) X_j w \, dz ds \\
&= \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} X_{j+n} \left\{ [X_j, X_{j+n}](v+sw) X_j w \right\} dz ds \\
&\qquad - \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} [X_j, X_{j+n}](v+sw) X_{j+n} X_j w \, dz ds \\
&\quad + \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=n+1}^{2n} X_{j-n} \left\{ [X_j, X_{j-n}](v+sw) X_j w \right\} dz ds \\
&\qquad - \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=n+1}^{2n} [X_j, X_{j-n}](v+sw) X_{j-n} X_j w \, dz ds \\
&= \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} X_{j+n} \left\{ -4\partial_t (v+sw) X_j w \right\} dz ds \\
&\qquad - \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} [X_j, X_{j+n}](v+sw) X_{j+n} X_j w \, dz ds \\
&\quad + \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=n+1}^{2n} X_{j-n} \left\{ 4\partial_t (v+sw) X_j w \right\} dz ds \\
&\qquad - \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=n+1}^{2n} [X_j, X_{j-n}](v+sw) X_{j-n} X_j w \, dz ds \\
&= \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} X_{j+n} \left\{ -4\partial_t (v+sw) X_j w \right\} dz ds \\
&\qquad - \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} [X_j, X_{j+n}](v+sw) X_{j+n} X_j w \, dz ds \\
&\quad + \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} X_j \left\{ 4\partial_t (v+sw) X_{n+j} w \right\} dz ds \\
&\qquad - \frac{3}{2} \int_0^1 \int_\Omega \sum_{j=1}^{n} [X_{j+n}, X_j](v+sw) X_j X_{j+n} w \, dz ds
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{3}{2}\int_0^1 \int_{\partial\Omega} \sum_{j=1}^{n} -4\partial_t(v+sw)X_j w\, \nu_{X_{j+n}}\, d\sigma(z)ds \\
&\qquad\qquad - \frac{3}{2}\int_0^1 \int_{\Omega} \sum_{j=1}^{n} [X_j, X_{j+n}](v+sw)X_{j+n}X_j w\, dzds \\
&\quad + \frac{3}{2}\int_0^1 \int_{\partial\Omega} \sum_{j=1}^{n} 4\partial_t(v+sw)X_{n+j} w\, \nu_{X_j}\, d\sigma(z)ds \\
&\qquad\qquad - \frac{3}{2}\int_0^1 \int_{\Omega} \sum_{j=1}^{n} [X_{j+n}, X_j](v+sw)X_j X_{j+n} w\, dzds \\
&= \frac{3}{2}\int_0^1 \int_{\partial\Omega} \sum_{j=1}^{n} -4\partial_t(v+sw)X_j w\, \nu_{X_{j+n}}\, d\sigma(z)ds \\
&\qquad\qquad - \frac{3}{2}\int_0^1 \int_{\Omega} \sum_{j=1}^{n} [X_j, X_{j+n}](v+sw)[X_{j+n}, X_j] w\, dzds \\
&\quad + \frac{3}{2}\int_0^1 \int_{\partial\Omega} \sum_{j=1}^{n} 4\partial_t(v+sw)X_{n+j} w\, \nu_{X_j}\, d\sigma(z)ds \\
&= -\frac{3}{2}\int_0^1 \int_{\Omega} \sum_{j=1}^{n} [X_j, X_{j+n}](v+sw)[X_{j+n}, X_j] w\, dzds \\
&= \frac{3n}{2}\int_0^1 \int_{\Omega} (4\partial_t)(v+sw)(4\partial_t) w\, dzds = 24n \int_0^1 \int_{\Omega} (\partial_t v + s\partial_t w)\partial_t w\, dzds \\
&= 12\, n \int_{\Omega} \{(\partial_t u)^2 - (\partial_t v)^2\}\, dz.
\end{aligned}
$$

This completes the proof of the first inequality of the theorem. The proof of the second one is similar.

□

4. O $\sigma_2(\mathcal{H})$–M

In this section we prove that if u is $\sigma_2(\mathcal{H})$–convex, we can locally control the integral of $\sigma_2(\mathcal{H})(u) + 12\, n\, (u_t)^2$ in terms of the oscillation of u. This estimate will be crucial for the L^2 estimate of $\partial_t u$.

Let us start with a lemma on $\sigma_2(\mathcal{H})$–convex functions.

Lemma 4.1. *If $u_1, u_2 \in C^2(\Omega)$ are $\sigma_2(\mathcal{H})$–convex, and f is convex in $\mathbb{R}^2$ and nondecreasing in each variable, then the composite function $w = f(u_1, u_2)$ is $\sigma_2(\mathcal{H})$–convex.*

Proof. Assume first that $f \in C^2(\mathbb{R}^2)$. We have

$$X_j w = \sum_{p=1}^{2} \frac{\partial f}{\partial u_p} X_j u_p,$$

$$X_iX_jw = \sum_{p=1}^{2}\left(\frac{\partial f}{\partial u_p}X_iX_ju_p + \sum_{q=1}^{2}\frac{\partial^2 f}{\partial u_q\partial u_p}X_iu_qX_ju_p\right),$$

and for every $h = (h_1, h_2) \in \mathbb{R}^2$

$$\begin{aligned}\langle\mathcal{H}(w)h, h\rangle &= \sum_{i,j=1}^{2n} X_iX_jw\, h_i\, h_j \\ &= \sum_{p=1}^{2}\frac{\partial f}{\partial u_p}\langle\mathcal{H}(u_p)h, h\rangle + \sum_{p,q=1}^{2}\frac{\partial^2 f}{\partial u_q\partial u_p}(\sum_{i=1}^{2n} X_iu_qh_i)(\sum_{j=1}^{2n} X_ju_ph_j).\end{aligned}$$

Since the trace and the second elementary symmetric function of the eigenvalues of the matrix $\mathcal{H}(u_p)$ are non negative, $\dfrac{\partial f}{\partial u_p} \geq 0$ for $p = 1, 2$, and the matrix

$$\left(\frac{\partial^2 f}{\partial u_q\partial u_p}\right)_{p,q=1,2}$$

is non negative definite, it follows that w is $\sigma_2(\mathcal{H})$–convex.

If f is only continuous, then given $h > 0$ let

$$f_h(x) = h^{-2}\int_{\mathbb{R}^2}\varphi\left(\frac{x-y}{h}\right)f(y)dy,$$

where $\varphi \in C^\infty$ is nonnegative vanishing outside the unit ball of $\mathbb{R}^2$, and $\int\varphi = 1$. Since f is convex, then f_h is convex and by the previous calculation $w_h = f_h(u_1, u_2)$ is $\sigma_2(\mathcal{H})$–convex. Since $w_h \to w$ uniformly on compact sets as $h \to 0$, we get that w is $\sigma_2(\mathcal{H})$–convex. □

Proposition 4.2. *Let $u \in C^2(\Omega)$ be $\sigma_2(\mathcal{H})$–convex. For any compact domain $\Omega' \Subset \Omega$ there exists a positive constant C depending on Ω' and Ω and independent of u, such that*

$$\int_{\Omega'}\{\sigma_2(\mathcal{H}(u)) + 12\, n\, (u_t)^2\}\, dz \leq C(\operatorname{osc}_\Omega u)^2. \tag{4.4}$$

Proof. Given $\xi_0 \in \Omega$ let $B_R = B_R(\xi_0)$ be a d–ball of radius R and center at ξ_0 such that $B_R \subset \Omega$. Let $B_{\sigma R}$ be the concentric ball of radius σR, with $0 < \sigma < 1$. Without loss of generality we can assume $\xi_0 = 0$, because the vector fields X_j are left invariant with respect to the group of translations. Let $M = \max_{B_R} u$, then $u - M \leq 0$ in B_R. Given $\varepsilon > 0$ we shall work with the function $u - M - \varepsilon < -\varepsilon$. In other words, by subtracting a constant, we may assume $u < -\varepsilon$ in B_R, for each given positive constant ε which will tend to zero at the end of the proof.

Define

$$m_0 = \inf_{B_R} u,$$

and

$$v(\xi) = \frac{m_0}{(1-\sigma^4)R^4}(R^4 - \|\xi\|^4).$$

Obviously $v = 0$ on ∂B_R and $v = m_0$ on $\partial B_{\sigma R}$. We claim that v is $\sigma_2(\mathcal{H})$–convex in B_R and $v \le m_0$ in $B_{\sigma R}$. Indeed, setting $r = \|\xi\|^4$, $h(r) = \dfrac{m_0}{(1-\sigma^4)R^4}(R^4 - r)$, and following the calculations in the proof of [4, Proposition 6.2] we get

$$\sigma_2(\mathcal{H}(v)) = c_n(|x|^2 + |y|^2)^2 \left(\frac{m_0}{(1-\sigma^4)R^4}\right)^2 \ge 0,$$

with c_n a positive constant and

$$\operatorname{trace}(\mathcal{H}(v)) = -(8n+4)\,(|x|^2 + |y|^2)\frac{m_0}{(1-\sigma^4)R^4} \ge 0,$$

because m_0 is negative. Hence v is $\sigma_2(\mathcal{H})$–convex in B_R. Since $v - m_0 = 0$ on $\partial B_{\sigma R}$, it follows from [4, Proposition 5.1] that $v \le m_0$ in $B_{\sigma R}$. In particular, $v \le u$ in $B_{\sigma R}$.

Let $\rho \in C_0^\infty(\mathbb{R}^2)$, radial with support in the Euclidean unit ball, $\int_{\mathbb{R}^2} \rho(x)\,dx = 1$, and let

$$f_h(x_1, x_2) = h^{-2} \int_{\mathbb{R}^2} \rho((x-y)/h)\, \max\{y_1, y_2\}\, dy_1 dy_2. \tag{4.5}$$

Define

$$w_h = f_h(u, v).$$

From Lemma 4.1 w_h is $\sigma_2(\mathcal{H})$–convex in B_R. If $y \in B_{\sigma R}$ then $v(y) \le u(y)$. If $v(y) < u(y)$ then $f_h(u,v)(y) = u(y)$ for h sufficiently small; and if $v(y) = u(y)$, then $f_h(u,v)(y) = u(y) + \alpha\, h$. Hence

$$\begin{aligned}\int_{B_{\sigma R}} \{\sigma_2(\mathcal{H}(u)) + 12\,n\,(\partial_t u)^2\}\,dz &= \int_{B_{\sigma R}} \{\sigma_2(\mathcal{H}(w_h)) + 12\,n\,((w_h)_t)^2\}\,dz \\ &\le \int_{B_R} \{\sigma_2(\mathcal{H}(w_h)) + 12\,n\,((w_h)_t)^2\}\,dz. \end{aligned} \tag{4.6}$$

Now notice that $f_h(u,v) \ge v$ in B_R for all h sufficiently small. In addition, $u < 0$ and $v = 0$ on ∂B_R so $f_h(u,v) = 0$ on ∂B_R. Then we can apply Theorem 3.1 to w_h and v to get

$$\begin{aligned}\int_{B_R} \{\sigma_2(\mathcal{H}(w_h)) + 12\,n\,(\partial_t w_h)^2\}\,dz &\le \int_{B_R} \{\sigma_2(\mathcal{H}(v)) + 12\,n\,(v_t)^2\}\,dz \\ &= \left(\frac{m_0}{(1-\sigma)R^4}\right)^2 \int_{B_R} (c_n(|x|^2 + |y|^2)^2 + 48n\,t^2)\,dz \\ &= \left(\frac{m_0}{(1-\sigma)}\right)^2 R^{2n-2} \int_{B_1} (c_n(|x|^2 + |y|^2)^2 + 48n\,t^2)\,dz. \end{aligned}$$

Combining this inequality with (4.6) we get

$$\int_{B_{\sigma R}} \{\sigma_2(\mathcal{H}(u)) + 12\,n\,(\partial_t u)^2\}\,dz \le C\,(m_0)^2 R^{2n-2} \le C\,R^{2n-2}(\operatorname{osc}_{B_R} u + \varepsilon)^2,$$

and then (4.4) follows letting $\varepsilon \to 0$ and covering Ω' with balls. $\square$

Corollary 4.3. *Let $u \in C^2(\Omega)$ be $\sigma_2(\mathcal{H})$–convex. For any compact domain $\Omega' \Subset \Omega$ there exists a positive constant C, independent of u, such that*

$$\int_{\Omega'} \sigma_2(\mathcal{H}(u))\, dz \leq C(\mathrm{osc}_\Omega u)^2 R^{2n-2}, \tag{4.7}$$

and

$$\int_{\Omega'} (\partial_t u)^2\, dz \leq C(\mathrm{osc}_\Omega u)^2 R^{2n-2}. \tag{4.8}$$

Corollary 4.4. *Let $u \in C^2(\Omega)$ be $\sigma_2(\mathcal{H})$–convex. For any compact domain $\Omega' \Subset \Omega$ there exists a positive constant C, independent of u, such that*

$$\int_{\Omega'} \mathrm{trace}\, \mathcal{H}_2(u)\, dz \leq C R^{2n} \mathrm{osc}_\Omega u. \tag{4.9}$$

4.1. **Measure generated by a $\sigma_2(\mathcal{H})$–convex function.** We shall prove that the notion $\int \sigma_2(\mathcal{H}(u)) + 12\, n\, u_t^2$ can be extended for continuous and $\sigma_2(\mathcal{H})$–convex functions as a Borel measure. We call this measure the $\sigma_2(\mathcal{H})$–measure associated with u, and we shall show that the map $u \in C(\Omega) \to \mu(u)$ is weakly continuous on $C(\Omega)$.

Theorem 4.5. *Given a $\sigma_2(\mathcal{H})$–convex function $u \in C(\Omega)$, there exists a unique Borel measure $\mu(u)$ such that when $u \in C^2(\Omega)$ we have*

$$\mu(u)(E) = \int_E \{\sigma_2(\mathcal{H}(u)) + 12\, n\, u_t^2\}\, dz \tag{4.10}$$

for any Borel set $E \subset \Omega$. Moreover, if $u_k \in C(\Omega)$ are $\sigma_2(\mathcal{H})$–convex, and $u_k \to u$ on compact subsets of Ω, then $\mu(u_k)$ converges weakly to $\mu(u)$, that is,

$$\int_\Omega f\, d\mu(u_k) \to \int_\Omega f\, d\mu(u), \tag{4.11}$$

for any $f \in C(\Omega)$ with compact support in Ω.

Proof. Let $u \in C(\Omega)$ be $\sigma_2(\mathcal{H})$–convex, and let $\{u_k\} \subset C^2(\Omega)$ be a sequence of $\sigma_2(\mathcal{H})$–convex functions converging to u uniformly on compacts of Ω. By Proposition 4.2

$$\int_{\Omega'} \{\sigma_2(\mathcal{H}(u_k)) + 12\, n\, (\partial_t u_k)^2\}\, dz$$

are uniformly bounded, for every $\Omega' \Subset \Omega$, and hence a subsequence of $(\sigma_2(\mathcal{H}(u_k)) + 12n(\partial_t u_k)^2)$ converges weakly in the sense of measures to a Borel measure $\mu(u)$ on Ω. Moreover, by the same argument used in the proof of [4, Theorem 6.5] the map $u \in C(\Omega) \to \mu(u) \in M(\Omega)$, the space of finite Borel measures on Ω, is well defined.

To prove (4.11), we first claim that it holds when $u_k \in C^2(\Omega)$. Indeed, let u_{k_m} be an arbitrary subsequence of u_k, so $u_{k_m} \to u$ locally uniformly as $m \to \infty$. By definition of $\mu(u)$, there is a subsequence $u_{k_{m_j}}$ such that $\mu\left(u_{k_{m_j}}\right) \to \mu(u)$ weakly as $j \to \infty$. Therefore, given $f \in C_0(\Omega)$, the sequence $\int_\Omega f\, d\mu(u_k)$ and an arbitrary subsequence $\int_\Omega f\, d\mu(u_{k_m})$, there exists a subsequence $\int_\Omega f\, d\mu(u_{k_{m_j}})$ converging to $\int_\Omega f\, d\mu(u)$ as $j \to \infty$ and (4.11) follows. For the general case, given k take $u_j^k \in C^2(\Omega)$ such that $u_j^k \to u_k$ locally uniformly as $j \to \infty$, and then argue as in the proof of [4, Theorem 6.5]. □

Corollary 4.6. *If $u \in C(\bar{\Omega})$ is $\sigma_2(\mathcal{H})$–convex in Ω, then $\mu(u)$ is a Radon measure.*

Proof. The measure $\mu(u)$ is Borel regular, see [3, p. 4-5] for definitions. Indeed, given $A \subset \mathbb{R}^{2n+1}$ there exists open sets V_k such that $A \subset V_k$ and $\mu(u)(V_k) \leq \mu(u)(A) + 1/k$ for all k. Thus, $\mu(u)(A) = \mu(u)(\cap_1^\infty V_k)$. Finally, the estimate (4.4) implies that $\mu(u)(K) < \infty$ for all compact K. Hence, $\mu(u)$ is a Radon measure. □

Corollary 4.7. *If $u, v \in C(\bar{\Omega})$ are $\sigma_2(\mathcal{H})$–convex in Ω, $u = v$ on $\partial\Omega$ and $u \geq v$ in Ω, then $\mu(u)(\Omega) \leq \mu(v)(\Omega)$.*

By arguing as in [4, Theorem 6.7] we also get the following comparison principle for $\sigma_2(\mathcal{H})$–measures.

Theorem 4.8. *Let $\Omega \subset \mathbb{R}^3$ be an open bounded set. If $u, v \in C(\bar{\Omega})$ are $\sigma_2(\mathcal{H})$–convex in Ω, $u \leq v$ on $\partial\Omega$ and $\mu(u)(E) \geq \mu(v)(E)$ for each $E \subset \Omega$ Borel set, then $u \leq v$ in Ω.*

5. A - $\mathcal{H}$–

As an application of our previous results we finally have the following main theorem.

Theorem 5.1. *If u is $\mathcal{H}$–convex, then $u \in BV^2_{\mathbb{H}}$ and so the distributional derivatives X_iX_ju exist a.e. for every $i, j = 1, \dots, 2n$.*

Proof. If u is $\mathcal{H}$–convex, then by [5, Theorem 3.1] u is locally Lipschitz continuous with respect to the distance d defined in (2.2), and X_iu exists a.e. for $i = 1, \dots, 2n$. Moreover, by [5, Theorem 4.2] there is a Radon measure $d\nu^{ij}$ such that, in the sense of distributions

$$\frac{X_iX_ju + X_jX_iu}{2} = d\nu^{ij}, \quad i, j = 1, \dots, 2n.$$

On the other hand, since u is continuous and $\sigma_2(\mathcal{H})$–convex, then by (4.8) $\partial_t u$ is in L^2_{loc}. Let $K \Subset \Omega$, $\phi = \sum_{j=1}^{2n} \phi_j X_j \in C^2(\Omega, \mathbb{R}^{2n+1})$, with compact support in K, $\|\phi\| < 1$. Since

$$X_iX_j = \frac{X_iX_j + X_jX_i + [X_i, X_j]}{2} = \frac{X_iX_j + X_jX_i}{2} \pm 2\delta_{i,i\mp n}\partial_t,$$

then for any $i = 1, \dots, 2n$

$$\begin{aligned}
\int_\Omega X_iu \operatorname{div}_X(\phi)dz &= -\int_\Omega u\, X_i \operatorname{div}_X(\phi)dz \\
&= -\sum_{j=1}^{2n} \int_\Omega uX_iX_j\phi_j dz \\
&= -\sum_{j=1}^{2n} \int_\Omega u\left(\frac{X_iX_j\phi_j + X_jX_i\phi_j}{2} \pm 2\delta_{i,i\mp n}\partial_t\phi_j\right) dz \qquad (5.12)\\
&= \sum_{j=1}^{2n} \int_\Omega \phi_j d\nu^{ij} \mp 2\sum_{j=1}^{2n} \delta_{j\mp n,j} \int_\Omega u\, \partial_t\phi_j dz \\
&\leq \sum_{j=1}^{2n} \nu^{ij}(K) \mp 2\sum_{j=1}^{2n} \delta_{j\mp n,j} \int_\Omega u\, \partial_t\phi_j dz.
\end{aligned}$$

Now, let u_ε be the horizontal mollification of the function u as in the proof of [5, Theorem 4.2]. Then u_ε is $\mathcal{H}$–convex and

$$\left|\int_\Omega u_\varepsilon \partial_t \phi_j dz\right| = \left|\int_\Omega \partial_t u_\varepsilon \phi_j dz\right| \leq c\|\partial_t u_\varepsilon\|_{L^2(K)} \leq C,$$

where c, C are positive constants depending on the diameter of K and on the oscillation of u over K, but independent of ε. Letting ε tend to zero, we get

$$\left|\int_\Omega u \partial_t \phi_j dz\right| \leq C. \tag{5.13}$$

Thus, by (5.12) and (5.13) we can conclude that

$$\int_\Omega X_i u \operatorname{div}_X(\phi) dz \leq \sum_{j=1}^{2n} \nu^{ij}(K) + C < \infty.$$

Hence, $u \in BV^2_{\mathbb{H}}$ and the result then follows from Theorem 2.3. □

6. A

Let A be an $n \times n$ symmetric matrix with eigenvalues $\lambda_1, \cdots, \lambda_n$, and the second elementary symmetric function

$$\sigma_2(A) = s(\lambda) = \sum_{j<k} \lambda_j \lambda_k$$

with $\lambda = (\lambda_1, \ldots, \lambda_n)$. An easy calculation shows that

$$\frac{\partial s}{\partial \lambda_j}(\lambda) = \sum_{k \neq j} \lambda_k$$

and

$$s(\lambda) = \frac{1}{2}\left\{\left(\sum_{j=1}^n \lambda_j\right)^2 - \sum_{j=1}^n \lambda_j^2\right\}. \tag{6.14}$$

Lemma 6.1. *If* $\sigma_2(A) \geq 0$ *and* $\operatorname{trace}(A) \geq 0$*, then* $\frac{\partial s}{\partial \lambda_j}(\lambda) \geq 0$ *for every* $j = 1, \ldots, n$*.*

Proof. Since

$$\operatorname{trace}(A) = \frac{\partial s}{\partial \lambda_j}(\lambda) + \lambda_j \geq 0,$$

then either $\lambda_j \geq 0$ or $\frac{\partial s}{\partial \lambda_j}(\lambda) \geq 0$. If $\lambda_j \geq 0$, since $s(\lambda) \geq 0$, then by (6.14)

$$\sum_{k=1}^n \lambda_k \geq \left(\sum_{k=1}^n \lambda_k^2\right)^{1/2} \geq \lambda_j,$$

and we get

$$\frac{\partial s}{\partial \lambda_j}(\lambda) = \sum_{k \neq j} \lambda_k = \sum_{k=1}^n \lambda_k - \lambda_j \geq 0.$$

□

Proposition 6.2. *If* $\sigma_2(A) \geq 0$ *and* $\mathrm{trace}(A) \geq 0$*, then*

$$\sum_{i,j=1}^{n} \frac{\partial \sigma_2}{\partial a_{ij}}(A) x_i x_j \geq 0$$

for every $x \in \mathbb{R}^n$.

Proof. Let C be a non negative definite Hermitian matrix. We write

$$\sigma_2(A + C) - \sigma_2(A) = s(\eta_1, \dots, \eta_n) - s(\lambda_1, \dots, \lambda_n)$$

where $\eta_1, \dots, \eta_n$ are the eigenvalues of $A + C$. Since $C \geq 0$, then $\eta_j \geq \lambda_j$, for any $j \in \{1, \dots, n\}$. Moreover, by Lemma 6.1, $\delta = \delta(A) = \dfrac{1}{2} \min \left\{ \dfrac{\partial s}{\partial \lambda_j}(\lambda_1, \dots, \lambda_n) : \ j = 1, \dots, n \right\} \geq 0$. If C is small enough, then

$$\begin{aligned}
\sigma_2(A + C) - \sigma_2(A) &= \int_0^1 \frac{d}{d\tau} s(\lambda + \tau(\eta - \lambda))\, d\tau \\
&= \sum_{j=1}^{n} \int_0^1 \frac{\partial s}{\partial \lambda_j}(\lambda + \tau(\eta - \lambda))\, d\tau\, (\eta_j - \lambda_j) \\
&\geq \delta \sum_{j=1}^{n} (\eta_j - \lambda_j) = \delta\, (\mathrm{trace}(A + C) - \mathrm{trace}(A)) \\
&= \delta\ \mathrm{trace}(C) \geq 0.
\end{aligned}$$

Let us now apply this inequality to the matrix

$$C = tx \cdot x^T = t(x_i x_j), \quad x \in \mathbb{R}^n,$$

and $t > 0$ small enough. We obtain

$$\sigma_2(A + tx \cdot x^T) - \sigma_2(A) \geq \delta\, \mathrm{trace}\, (C) = \delta t |x|^2. \tag{6.15}$$

On the other hand

$$\frac{d}{dt} \sigma_2(A + tx \cdot x^T)\ |_{t=0} = \sum_{i,j=1}^{n} \frac{\partial \sigma_2}{\partial a_{ij}}(A) x_i x_j.$$

Then, from (6.15) we get

$$\sum_{i,j=1}^{n} \frac{\partial \sigma_2}{\partial a_{ij}}(A) x_i x_j \geq \delta |x|^2 \geq 0, \quad \forall x \in \mathbb{R}^n. \tag{6.16}$$

□

R

[1] L. Ambrosio and V. Magnani, *Weak diferentiability of BV functions on stratified groups*. `http://cvgmt.sns.it/papers/ambmag02/`

[2] D. Danielli, N. Garofalo, and D. M. Nhieu, *Notions of convexity in Carnot groups*. Comm. in Analysis and Geometry, **11** 2 (2003) 263–341.

[3] L. C. Evans and R. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton 1992.

[4] C. E. Gutiérrez and A. Montanari, *Maximum and comparison principles on the Heisenberg group*. ArXiv preprint 2003.

[5] Guozhen Lu, J. Manfredi, and B. Stroffolini, *Convex functions on the Heisenberg group.* To appear in Calculus of Variations.
[6] E. M. Stein, *Harmonic Analysis: Real Variable methods, Orthogonality and Oscillatory Integrals.* Vol. 43 of the Princeton Math. Series. Princeton U. Press. Princeton, NJ, 1993.
[7] N. S. Trudinger and Xu-Jia Wang, *Hessian measures I.* Topol. Methods Nonlinear Anal. **10** (1997) 225-239.

D M , T U , P , PA 19122
E-mail address: gutierrez@math.temple.edu

D M , U ` B , P P S D 5, B , 40127, I
E-mail address: montanar@dm.unibo.it